ABSTRACT

The present invention is directed to converting text to speech such that a more natural sounding speech output is generated compared to most currently available text to speech engines. The invention does so in a computationally efficient manner that is suitable for supporting hundreds of channels on a single application server. It provides a vocabulary of words that covers over 95% of words typically found in e-mails, with the remaining words, names, etc. being covered by a second text to speech engine. The second text to speech engine can be a more computationally intensive speech synthesis engine without much impact to the overall computational efficiency of the text to speech system, since it only needs to handle the remaining 5% of the words. The invention can integrate the words generated by the second text to speech engine seamlessly with the words generated by the first engine. Another benefit of the invention is that creating new 'voices' for the text to speech engine is simple and inexpensive. Allowing voices to be created that match pre-recorded "voice prompts" in a voice messaging system, for example.